Preprint
UCRL-JC-142539

# On the Construction of the Prediction Error Covariance Matrix

*T. Waseda, L. Jameson, M. Yaremchuk, H. Mitsudera*

This article was submitted to ***The Oceanographic Society of Japan Spring Symposium, 2001, Tokyo, Japan, March 27-31, 2001***

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

**February 2, 2001**

This work was performed under the auspices of the United States Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

# On the construction of the prediction error covariance matrix

∘T. Waseda (IPRC,FRSGC), L. Jameson (LLNL), M. Yaremchuk (IPRC), H. Mitsudera (IPRC,FRSGC)

KEYWORD: Assimilation, Kalman Filter, EOF, Wavelet

**Abstract**

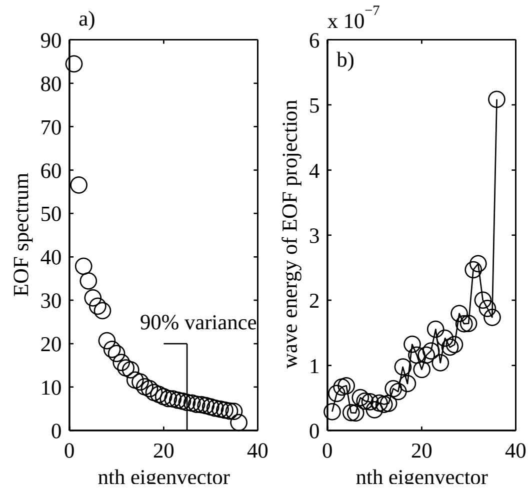
Implementation of a full Kalman filtering scheme in a large OGCM is unrealistic without simplification and one generally reduces the degrees of freedom of the system by prescribing the structure of the prediction error. However, reductions are often made without any objective measure of their appropriateness. In this report, we present results from an ongoing effort to best construct the prediction error capturing the essential ingredients of the system error that includes both a correlated (global) error and a relatively uncorrelated (local) error. The former will be captured by an EOF modes of the model variance whereas the latter can be detected by wavelet analysis.

**EOFs: identification of patterns in the error**

When the error associated with the estimation of the state vector, $\mathbf{x}$, of length $n$ has no time invariant pattern, the Kalman filtering scheme would require $O(2n)$ computation and $O(n^2)$ storage. Then, the application of the full Kalman filtering scheme is rather unrealistic for a typical OGCM of $n \sim O(10^7)$. However, it is in general, considered that there are some degree of correlation amongst the error of the elements of the state vector and so the error can be approximated reasonably well by a linear combination of a small set of those patterns (e.g. SEEK filter, Veron *et al.* *JGR* **1999**). These patterns can be found as eigenvectors of the error covariance matrix. Because the true state vector is unknown, in reality, one would typically replace the true state vector with the model mean and an ensemble averaging will be replaced by time averaging. Under these assumptions, we can make a first guess of the structure of the prediction error by constructing an EOF modes of the model data; in this study, we analyzed the SSH field from the regional OGCM. The resulting spectrum shown in figure a shows a sharp dropoff of the spectrum indicating that only $r = O(10)$ modes are needed to approximate the prediction error; this is a huge reduction in the degrees of freedom.

**Wavelets: detection of the numerical error**

Any computational data are subject to numerical errors associated with their deviation from low-order polynomial structure, and the relative magnitude of those can be detected by wavelet analysis (Jameson, *J. Sci. Comput.*, **1999**). These relatively uncorrelated errors are local but would accumulate over time to cause global error in the model prediction. The knowledge of the numerical error was utilized to model the prediction error covariance matrix in a simple setting and with that the assimilation skill improved over the traditional optimal interpolation assimilation (EEWADAi, Jameson and Waseda, *J. Atmos. Ocean. Tech.* **2000**). The additional computational resource required to conduct the wavelet diagnostic is moderate. Furthermore, it is straighforward to increase the scale of the wavelets to detect small scale variation of the model data that are not necessarily related with the numerical error but are too small to be represented by EOFs.




**Orthogonality of the EOFs and the Wavelets**

Whether the global error or the local error are more suitable to represent the model error depends on the nature of the data. In oceanographic applications, the data is primarily composed of large scale slowly varying flows with localized small-scale phenomenon appearing periodically. Thus, the best approximation of the prediction error would likely involve both global and local error. In order to do so, their independence should be tested. We have thus projected the wavelet bases onto the EOF modes, figure b. It is readily observed that the scale of the wavelets matches more for the higher EOF modes that has finer scale. This result indicates that a rather simple linear combination of the global error and the local error is possible.

**Combining the EOFs and Wavelets**

We can thus approximate the prediction error as:

$$\mathbf{P}^f \approx \tilde{\mathbf{L}}\tilde{\mathbf{\Lambda}}\tilde{\mathbf{L}}^T + \mathbf{Q}, \qquad (1)$$

where $\tilde{\mathbf{L}}$ contain the EOF eigenvectors, the corresponding amplitudes are contained in matrix $\tilde{\mathbf{\Lambda}}$ and $\mathbf{Q}$ contains the magnitudes of the wavelet projection onto the data. While the slowly varying $\tilde{\mathbf{\Lambda}}$ can be prognostically computed, the rapidly changing $\mathbf{Q}$ should be diagnostically detected.

**Conclusion**

We have presented a possibility of constructing the prediction error by a linear combination of two types of errors of different origin. The correlated global error will be modeled by $r$ EOF modes, whereas the uncorrelated local error will be diagnostically detected using wavelet analysis.

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.